Applied Soft Computing 38 (2016) 754–761

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A chaotic time series prediction model for speech signal encoding based on genetic programming

Lei Yang[c], Junxi Zhang[d], Xiaojun Wu[a,b,c,*], Yumei Zhang[b,c], Jingjing Li[a]

[a] Key Laboratory of Modern Teaching Technology, Ministry of Education (Shaanxi Normal University), Xi'an 710062, China
[b] School of Computer Science, Shaanxi Normal University, Xi'an 710062, China
[c] School of Automation, Northwestern Polytechnical University, Xi'an 710072, China
[d] Department of Vehicle Engineering, Xi'an Aeronautical, Xi'an 710077, China

ARTICLE INFO

Article history:
Received 22 August 2014
Received in revised form 28 August 2015
Accepted 3 October 2015
Available online 20 October 2015

Keywords:
Chaotic time series prediction
Genetic programming
Nonlinear coding model

ABSTRACT

In this paper, a novel solving method for speech signal chaotic time series prediction model was proposed. A phase space was reconstructed based on speech signal's chaotic characteristics and the genetic programming (GP) algorithm was introduced for solving the speech chaotic time series prediction models on the phase space with the embedding dimension m and time delay τ. And then, the speech signal's chaotic time series models were built. By standardized processing of these models and optimizing parameters, a speech signal's coding model of chaotic time series with certain generalization ability was obtained. At last, the experimental results showed that the proposed method can get the speech signal chaotic time series prediction models much more effectively, and had a better coding accuracy than linear predictive coding (LPC) algorithms and neural network model.

1. Introduction

The Linear Prediction Coding (LPC) has been widely used in speech signal processing field. It is based on the correlation between the speech samples, using a linear model, with the past p sample values to predict the sample values present or future. Then, a large number of studies on non-linear characteristics of speech signal found that speech signal and the discrete speech signal sequence showed a very complicated nonlinear process and obvious chaotic characteristics [1–4]. Nonlinear science inspired the study of speech signal which can be analyzed by non-linear methods. The research of nonlinear models for speech signal coding became a hot spot of speech signal processing, and using chaos theory to research speech signal has made some achievements [4–6]. It is proved that the reconstructed speech signal helped to analyze its characteristics deeply and accurately [6]. Therefore, the introduction of chaotic time series analysis theory has very important theoretical and practical values for not only researching the chaos characteristics and the processing methods of speech signal, but also constructing nonlinear speech signal processing models.

The neural network model is the most commonly used method for nonlinear speech signal processing, and has some achievements [7–9]. It has been applied to speech recognition [10], speech transmission [11], etc.

Thyssen et al. [12] hold that the speech signal was the combination of linear and nonlinear function. Linear prediction method has been used for modeling the linear part. They proved that the residual signal has obvious nonlinear characteristics, and used Volterra and neural network methods to modeling the residual part which is nonlinear part of the signal. The combination of those linear models and nonlinear models has improved the accuracy of coding. Atthew K. Luka established a local ethnic languages recognition system through the neural network technology [13]. This system uses Mel-frequency Cepstral Coefficients of the speech signal that extract from Mel filter banks as the input of the multilayer neural network. They combined the conjugate gradient backward propagation algorithm with neural network to improving the convergence rate, which shows good recognition ability. Using chaos theory to analyze speech signal, Lee and Tong [14] and Aina et al. [15] applied the traditional local linear and radial neural network to the analysis of chaotic time sequence and achieved two kinds of nonlinear speech coding predictors. The result showed that the predictor established in the reconstructed phase space had significantly improved compared with the linear predictor.

The neural network model has been successfully used in the speech signal coding and also provides a structured prediction model. In recent years, it has made some achievements in the nonlinear prediction of the speech signal sequence. But the structure of the neural network model is based on the optimization of the weights; it is not conducive to the speech signal's analysis and processing as it cannot provide the fixed structure for the models to describe the speech signal like LPC, and it needs to rebuild for different samples. So its practical application has been limited. To solve these problems, this paper introduces GP (Genetic Programming) as solving method for the speech signal prediction models. Compared to neural network method, GP can obtain the explicit model structure which can facilitate its analysis and application. As early as 1998, Conrads et al. [16] used GP in speech; they found that GP could find programs to discriminate certain spoken vowels and consonants. Xie and Zhang [17,18] did a series of research in rhythmic stress detection in spoken English using GP. Later, GP was proposed to construct nonlinear speech coding models [19].

GP is a special optimization algorithm developing from GA (Genetic Algorithm) [20]. It uses hierarchical structure to describe programming problems. Issues in different fields are attributed to computer program that searches for solutions to meet fixed constraint conditions. GP algorithm can take advantage of limited samples to discover parallel estimate of the model structure and its parameters. In recent years,

* Corresponding author. Tel.: +86 13991161772.
E-mail address: xjwu@snnu.edu.cn (X. Wu).

http://dx.doi.org/10.1016/j.asoc.2015.10.003
1568-4946/

it has been widely used in nonlinear system modeling [21], data analysis [22] and other areas.

Based on the following four reasons GP can be used to solve the problem of speech signal processing. Firstly, compared to LPC, GP highlights the non-linear characteristics. Secondly, compared to the Artificial Neural Network, it can obtain the explicit model structure, which is facilitated for its analysis, optimization and application. Thirdly, compared to other non-linear models, using GP models, we can achieve the overall modeling without considering the linear or non-linear part. Finally, compared to the Volterra model, it overcame the weakly nonlinear problem of the models, and it is more streamlined and more efficient.

Considering chaotic characteristics of the speech signal, this paper provides an improved genetic programming algorithm to establish the non-linear model of the speech signal; through the model analysis, a model structure which has generalization ability can be chosen. By optimizing parameters, speech signal encoding could be achieved and it provides a new model for speech signal's nonlinear analysis.

2. GP algorithm for nonlinear modeling

Compared to other algorithms, the GP algorithm can better gain a nonlinear model with minimum error, and it does not need any prior knowledge. The evolutionary process changes the structure of the model automatically through crossover and mutation, calculates the fitness of new individuals by the constructed fitness function and keeps the optimal one, to make the evolution direction toward the better target. As the model that is produced by GP algorithm has explicit structure, it can intuitively describe the problem to be solved to a large extent, so it is widely used in the problem of nonlinear modeling, especially in time series prediction model [14,23].

Lee and Tong separated complete time series into linear and nonlinear parts according to the analysis of the time series' composition, built the prediction model of linear part using ARIMI algorithm and a nonlinear model with the GP algorithm for the residual sequence [14]. The nonlinear model that describes the time series is obtained and it is effectively improved than linear method. It benefits the combination of these two kinds of models. Wagner et al. separated the time series environment into static environment and dynamic environment, and applied the GP algorithm to the dynamic environment (DyFor GP), because they believe that the traditional GP algorithm does not consider the dynamic factors in the environment when predicting time series [22]. In the evolutionary process, the method changes and adjusts the order dynamically depending on the different environments to make the model be more optimized. Estevez et al. [24] researched the identification of the speech signal and noise signal using GP, the function set contains some logical operators such as AND, OR, NOT, GT (greater-than). The composition of Terminator set is the constant R and the terminator defined in ITU-T G.729B VAD standard. Finally, the judgment method of speech signal was presented. Compared to standard G.729B, the judgment of speech signal and non-speech signal by using the method is more accurate.

GP algorithm solves the nonlinear problems effectively, and its explicit model structure is more conducive for data analyzing and processing. This paper works on the speech signal coding model with GP for getting more valuable model structure.

3. Proposed speech coding method

The speech signal is a chaotic system that produces orders and laws from disorder and complexity. The chaotic attractor is one of the chaotic characteristics indicating the regularity of a chaotic system. The inherent regularity of the chaotic system shows that it is predictable. This paper uses chaos system's internal regularity, combined with nonlinear prediction method (GP) to predict and build models of the speech signal.

In this paper, the speech signal processing flow chart, as shown in Fig. 1, includes five contents:

(1) Process the speech signal, enhance high-frequency part of the speech signal through the pre-emphasis, and then use the phase space reconstruction technique to reconstruct the discrete sequence.
(2) According to the characteristics of the nonlinear model of the speech signal to improve the GP algorithm, and solve the "best" non-linear model of each frame speech signal by using the improved GP algorithm.
(3) Summarize and analyze the prediction model in the different samples, and get a model structure with good generalization ability.
(4) On the basis of model structure, use DUPSO optimization algorithm to solve the prediction parameters of each frame of data.
(5) Evaluate the coding model in this article by experiments.

On the basis of the speech signal's chaotic characteristics, this paper provides a method for solving speech signal's nonlinear model by improved GP algorithm. As a result of the limitations of samples used in the modeling process, the model structure may only apply to the limited features speech samples. Other speech samples may have different characteristics which need different models to describe their structure, this study tries to make the model for the sample has adaptive capacity (generalization ability). But more samples still need a different classification for models. Therefore, it is a problem which needs further study to build classified model for speech sample characteristics, and needs more scientists to participate in. With the development of research, better description for the nonlinear model of speech signal will be found and applied.

4. Prediction model of speech signal based on improved GP algorithm

As using the traditional linear prediction model for speech coding, the model structure is fixed. The encoding and decoding of speech signal can only transmit different parameters which will be decoded at the receiver.

Wu and Yang [19] make improvements in two ways on the GP algorithm. First, in the initialization of the population, a variety of groups are used, in order to increase the diversity of solutions and improve the global search capability. Second, the hill-climbing algorithm is introduced for further optimizing model parameters which are based on structural optimization, so the GP algorithm gets more objective evaluation of the model structure in the evolutionary process.

In this paper, after the speech signal pre-processing and reconstruction phase space, we use that improved GP algorithm to modeling each frame data, and then analyze the structure of the model, select one or a group of standardized models with better generalization ability. Finally, optimization algorithm is used to

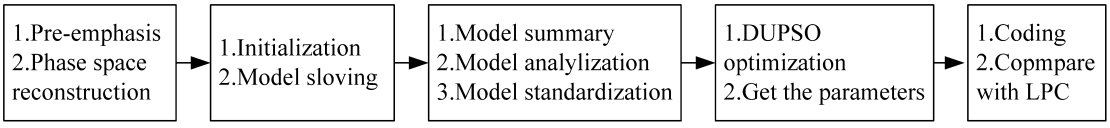

Fig. 1. Structure diagram of this research.